Journal of Demographic Economics (2021), **87**, 479–510
doi:10.1017/dem.2020.14

CAMBRIDGE
UNIVERSITY PRESS

RESEARCH PAPER

The impact of natural disasters on migration: findings from Vietnam

Trong-Anh Trinh*, Simon Feeny and Alberto Posso

School of Economics, Finance and Marketing, RMIT University, Melbourne, VIC 3000, Australia
*Corresponding author. E-mail: tronganh.trinh@rmit.edu.au

(Received 23 April 2019; revised 13 May 2020; accepted 13 May 2020)

Abstract

Increasingly, studies are examining whether the incidence of natural disasters influences household migration. This paper examines whether the severity of natural disasters is important for migration decisions in Vietnam, rather than just examining their occurrence. Data for a sample of 1,003 farm households from the Vietnam Household Living Standard Survey are examined for the period 2006–2008. A residual generated regressor approach is adopted to isolate the direct impact of disasters on migration from the indirect impact they have on migration through reducing agricultural output and income. Findings suggest that more severe disasters are directly associated with a greater probability of migration. Furthermore, such outcomes are the same for poor households *vis-à-vis* their non-poor counterparts.

Key words: Agriculture; migration; natural disasters; residual generated regressor; Vietnam

JEL classification: J61; O15; P25; Q54

1. Introduction

Migration is inherently a spatial phenomenon which represents the change in an individual's usual place of residence to another area over a given period of time. Migration is conventionally driven by various economic, political, and social determinants, as well as demographic characteristics. Recent views hold that climate change, especially through an increase in the frequency and severity of natural disasters, is expected to bring about significant changes in migration patterns for communities both within and between countries. According to a special report by the Asian Development Bank, more than 42 million people were displaced in the Asia Pacific region during 2010 and 2011 due to natural disasters [Édes *et al.* (2012)]. The fifth assessment report of the Intergovernmental Panel on Climate Change confirms that one of the gravest effects of extreme climatic events in the future will be that of human migration [Pachauri *et al.* (2014)].

This paper uses two waves of household data (2006 and 2008) to examine both the direct and indirect impacts of disasters on migration in Vietnam; a country that has in recent years' undergone fast-paced urbanization resulting from a high rate of industrialization. However, it remains predominantly agricultural and highly vulnerable to weather shocks. The paper departs from much of the existing literature by adopting a residual generated regressor approach and by examining the role of disaster intensity rather than disaster occurrence. Findings from a range of empirical models suggest that more severe disasters are associated with a greater probability of migration.

The remainder of this paper is structured as follows. Section 2 provides a background to the paper, summarizing the approach and findings of previous studies as well as outlining the paper's contributions. Section 3 outlines the vulnerability of Vietnam to natural hazards and provides the motivation for focusing on this particular country. Section 4 provides the empirical model and methodological approach used to estimate the relationship between weather shocks and migration. This section also describes the data. Section 5 discusses the empirical findings and section 6 provides implications from the paper's findings.

2. Background and contribution

There is a large body of literature that examines migration as a coping strategy to natural disasters. This literature is implicitly motivated by the theoretical model of migration suggested by Harris and Todaro (1970) and further developed by Stark and Bloom (1985). A main assumption is that households exposed to risk consider migration as a mechanism to diversify income. Since inter-household family transfers are an important source of insurance in low-income countries, a household, as long as it can afford the cost of migration, may decide to relocate one or more members to reduce the gap of income shocks between origin and destination [Rosenzweig and Stark (1989)].

Given this theoretical background, a growing number of studies have provided empirical evidence of migration induced by these types of shocks.[1] Following the seminal work by Rosenzweig and Stark (1989) in India, the impact of weather shocks on migration has been found in both low-income [Findley (1994), Meze-Hausken (2000), Ezra and Kiros (2001), Henry *et al.* (2004)] and high-income countries [McLeman and Smit (2006), Feng *et al.* (2010)].[2] However, the size of this impact is not clear with migration being a highly contextual phenomenon, depending on political, economic, and social factors. For example, Barrios *et al.* (2006) show that rainfall shocks increase rural–urban migration in sub-Saharan Africa, whereas Marchiori *et al.* (2012) show that migration in sub-Saharan Africa can be constrained by access to credit, available technologies, and macroeconomic conditions. This suggests that further empirical evidence of the nexus between climatic events and migration is warranted, particularly in societies that are vulnerable to climatic events and which are heavily reliant on agriculture for their livelihood.

Although most studies focus on internal migration [e.g., Bohra-Mishra *et al.* (2014), Bohra-Mishra *et al.* (2017), Dallmann and Millock (2017), Peri and Sasahara (2019)], a

[1]Income shocks do not necessarily increase the probability of migration from a theoretical point of view. Clemens (2014) provides a summary of theoretical models that explain the relationship between economic development and international migration. He finds an inverted-U relationship whereby emigration from developing countries generally rises with development but starts to fall when countries reach upper-middle income status. Clemens and Postel (2018) find that foreign aid has played little role in deterring emigration from developing countries. Gray and Mueller (2012) demonstrate that the propensity to migrate depends on the barriers faced by vulnerable households as well as their adaptive capacity in the context of Bangladesh.

[2]For a recent review of the impact of natural disasters on migration, see Thiede and Gray (2017) and Berlemann and Steinhardt (2017).

number of recent studies have examined international migration induced by shocks. For example, Baez *et al.* (2017a, 2017b) investigated youth migration caused by droughts and heat exposure in Latin America and the Caribbean. The authors find a significant impact of shocks, although this impact is mediated by social protection and government policies. Studying both natural disasters and long-term climatic factors, Beine and Parsons (2015) find little evidence of an association between the two, although a later study shows that natural disasters tend to increase emigration to neighboring countries [Beine and Parsons (2017)]. Other studies using cross-country data reach similar conclusions [Drabo and Mbaye (2015), Cattaneo and Peri (2016), Gröschl and Steinwachs (2017)].

The empirical approach adopted in most studies testing the relationship between natural disasters and migration also builds on the implicit assumption that the occurrence of an extreme event will reduce household income. Most studies employ two-stage least squares techniques, whereby the first stage examines the relationship between shocks (as an instrument) and household income, whereas the second stage estimates the impact of household income on the probability of migration. In doing so, two important issues need addressing. First, an extreme event can directly impact migration decisions by damaging or destroying household infrastructure. Natural disaster variables, therefore, should also be included in the second stage of these empirical models. However, this approach is problematic because of the expected correlation between natural disasters and household income. Multicollinearity can lead to the following symptoms: (i) coefficients may have very high standard errors; and (ii) coefficients may have the "wrong" sign or be of implausible magnitudes [see Greene (2012), p. 129]. In other words, parameter estimates can be unreliable and the standard errors attached to the parameters will be inflated. This presents an empirical challenge since including only one of these variables in the model will result in omitted variable bias.

The second issue relates to the measurement of extreme events or weather shocks.[3] Most studies to date employ long historical data on temperature and precipitation to measure weather shocks [Maccini and Yang (2009), Datar et al. (2013), Baez *et al.* (2017c)]. Specifically, the occurrence of an extreme event is captured by a binary variable, taking the value of one when temperature or precipitation deviates from an arbitrary distance from the long-term average. One limitation of this approach arises when a disaster occurs with trivial loss. Such a case is unlikely to lead to migration. Conversely, households might face little choice but to migrate in the advent of a shock of large magnitude. Consequently, there is an increasing number of studies using more detailed information on natural disasters to capture their intensity [e.g., Beine and Parsons (2017), Bohra-Mishra *et al.* (2017), Dallmann and Millock (2017)].

This study contributes to the literature by addressing these two issues. Regarding the empirical model, we adopt a novel two-stage econometric technique that accounts for the direct impact of natural disasters on migration as well as its potential indirect impact via reduced household income. The technique is known as the residual generated regressor approach.[4] First, agricultural output is regressed on the disaster variables with other controls following the empirical model of Mendelsohn *et al.*

[3]As in other studies, we use the terms "natural disasters" and "weather shocks" interchangeably [Gray and Mueller (2012), Gröger and Zylberberg (2016)].

[4]See Gomanee *et al.* (2005) for an example of how this technique has been applied in the foreign aid effectiveness literature.

https://doi.org/10.1017/dem.2020.14 Published online by Cambridge University Press

(1994). We use agricultural output as a proxy for income since we are interested in farmers living in rural areas where agriculture is the main source of income. Findings from the estimation of this model provide an indication of the extent that natural disasters have impacted on agricultural output. The residual from this model is then employed as an explanatory variable in a second model that estimates the probability of migration. Note that the residual captures the variation in agricultural output that is not explained by shocks and will not therefore be correlated with the shock variables. This approach: (i) allows for the reliable identification of the direct impact of shocks on migration; and (ii) controls for changes in agricultural output and income (not due to shocks but other factors) which can also impact on migration decisions.

We follow the recent literature by examining the *severity* of natural disasters by using the number of deaths, injuries, houses damaged and houses destroyed, rather than examining the impact of the *occurrence* of natural disasters. By employing the magnitude of disasters rather than their occurrence, we account for the possibility that shocks with higher intensity are more likely to trigger migration, whereas shocks causing less damage are less likely to lead to farm-household members relocating. Since migration is an economic decision, what matters is the way that disasters translate into tangible outcomes. This suggests that to properly quantify weather shocks it is useful to focus on outcome variables rather than on deviations in temperature or precipitation alone. That said, as a robustness check, we also employ binary shock variables defined as extreme deviations of temperature and precipitation.

We choose Vietnam as a case study to investigate the relationship between natural disasters and migration. The country is one of five countries deemed most-affected by climate change and suffers from a high frequency of natural disasters [ISPONRE (2009)]. These extreme events have posed a significant threat to large portions of the population living in rural areas with agricultural production being their main source of income. Given their vulnerability to disasters, affected farm-households are likely to consider migration as a possible coping strategy. The data used in this study are sourced from the Disaster Inventory System (DesInventar) provided by the United Nations Office for Disaster Risk Reduction (UNISDR). This dataset provides unique information on the damage from disasters at the province level. We then match this disaster data with household data in Vietnam by using the Vietnam Household Living Standard Survey (VHLSS) for the period 2006–2008. We test the hypothesis that a higher severity of natural disasters, measured by the four indicators identified above, is associated with a higher probability of migration.

We further explore the relationship between natural disasters and migration by looking at different household characteristics. Specifically, we assess separately poor and non-poor households using the internationally accepted poverty line of PPP $1.25-a-day (in 2008). This is motivated by contradictory findings from the recent literature. For example, some studies demonstrate that low-income people often settle in areas vulnerable to extreme climate events which force them to subsequently migrate [Morrow-Jones and Morrow-Jones (1991), Koerber (2006)]. Conversely, Myers *et al.* (2008) find that wealthier households are more likely to migrate from flood prone areas, since poorer families are unable to meet the costs of migration.

Finally, we contribute to the literature by examining the relationship between migration and natural disasters in the context of public support. Vietnam provides an interesting case study with different post-disaster support programs from the government and international organizations. For example, the government together with the World Bank has established a "Community-based disaster risk management program" in the Mekong Delta to build flood and storm mitigation infrastructures, aimed at strengthening the capacity of people in this region to cope with extreme events. This study seeks to evaluate the effectiveness of the program by testing whether households receiving benefits from the government program in Mekong Delta have had better coping abilities following a disaster and therefore a lower probability of migration.

3. Natural hazards in Vietnam

This paper provides a case study of Vietnam. Located in a tropical monsoon region combined with a diverse and complex topography, Vietnam has suffered different types of natural disasters: both hydro-meteorological (floods, storms, and droughts) and geophysical (landslides and earthquakes). Figure 1 summarizes the most common types of natural hazards that occurred during the period 1990–2010. According to Figure 1a, floods are the most frequent events accounting for 48% of disasters, followed by hailstorms (20%), storms (13%), and flash floods (7%). Cyclones (or typhoons), landslides, and other disasters account for 12% of the reported events. In terms of disaster damage, Figure 1b shows that floods account for 67% of deaths and floods, hailstorms, storms, and flash floods combined are responsible for nearly 90% of the loss of life from disasters [IMHEN and UNDP (2015)]. By region, Figure 2 shows that most affected provinces by natural disasters are located in the North and Central Vietnam. Provinces in the Mekong River Delta have a smaller number of reported events than the average. However, they suffer from a higher number of people affected and households damaged [IMHEN and UNDP (2015)].

Despite its vulnerability to shocks, there is little quantitative research examining disaster induced migration in Vietnam. One important exception is Gröger and Zylberberg (2016) who use panel data for 2,200 households in three provinces across 3 years (2007, 2008, and 2010). The authors find that households exposed to Typhoon Ketsana in 2009 experienced an average reduction in income of 10%, and around 17% of households responded by sending members away to diversify income. Another exception is Koubi *et al.* (2016) who examine different types of environmental stressors, namely, short-term and long-term events. The authors show that sudden-onset environmental events, such as floods and typhoons, are more likely to trigger migration, whereas long-term environmental problems, such as salinity, reduce the likelihood of migration. This suggests that individuals respond to longer-term events with other forms of adaptation, rather than migration.

This paper builds upon these previous studies by using national household surveys collected across Vietnam during the period 2006–2008. This study differs from Gröger and Zylberberg (2016) by examining the effects of all disasters occurring over this period instead of concentrating on specific events. Although Koubi *et al.* (2016) use the occurrence of disasters to examine its impact on migration, we account for the possibility that shocks with higher intensity will trigger migration, whereas shocks leading to less damage are less likely to lead to migration. Finally, this study also proposes a novel econometric framework to assess both the direct and indirect impacts of natural disasters on migration.

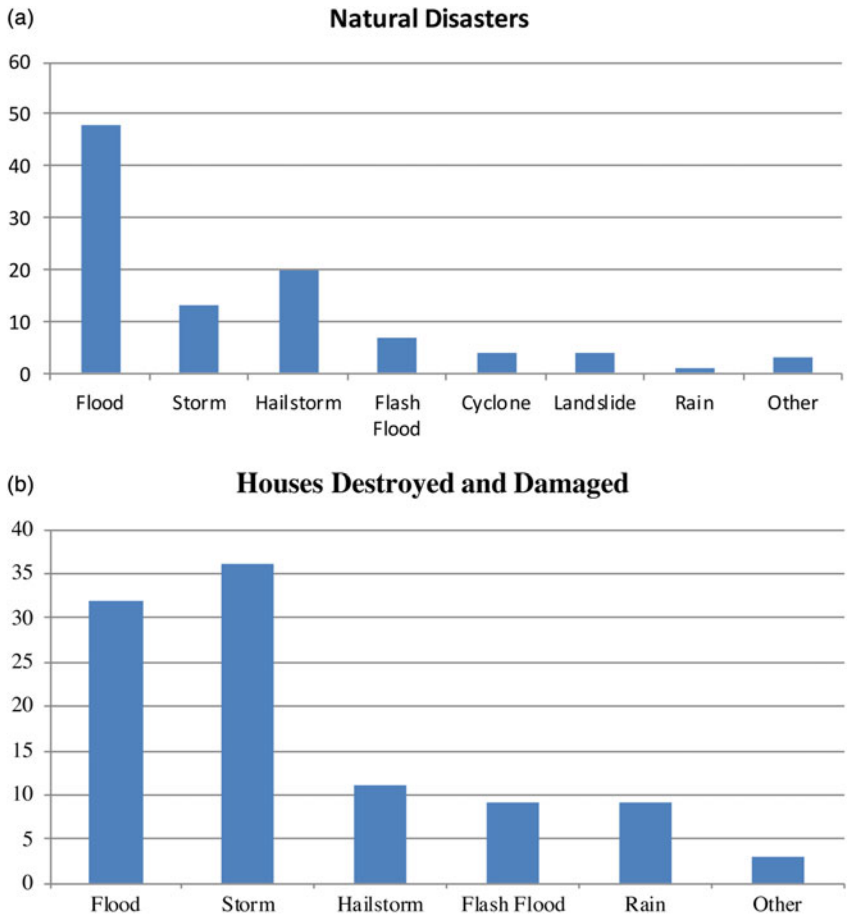

Figure 1. (a) Proportion of disasters (%) in terms of reported events in 2006–2008. (b) Proportion of disasters (%) in terms of damage in 2006–2008.
Notes: Flash flood is defined as an event that occurs within 6 hours following the end of the causative event, whereas flood is defined as an event that occurs after 6 hours following the end of the causative event.
Source: IMHEN and UNDP (2015).

4. Model specification and data

4.1 Model specification

We use a cross-sectional sample of households located in the agricultural sector to investigate the impact of natural disasters on the migration decisions of farm households in Vietnam. Following previous studies, we start with a basic framework by employing a traditional IV model to examine the relationship between weather shocks and migration decisions [Cai *et al.* (2016), Kubik and Maurel (2016)]. The first stage of this model examines the impact of natural disasters on agricultural output as presented in equation (1). The second stage then tests if a reduction in agricultural output caused by shocks leads to an increase in the probability of migration:

$$\text{Crop\_revenue}_{i,j,t-1} = \alpha_0 + \alpha_1' \text{Shock}_{j,t-1} + \alpha_2' X_{i,j,t} + \alpha_3 C_{i,j} + \alpha_4 E_{i,j,t} + \upsilon_{i,j,t-1} \quad (1)$$

https://doi.org/10.1017/dem.2020.14 Published online by Cambridge University Press

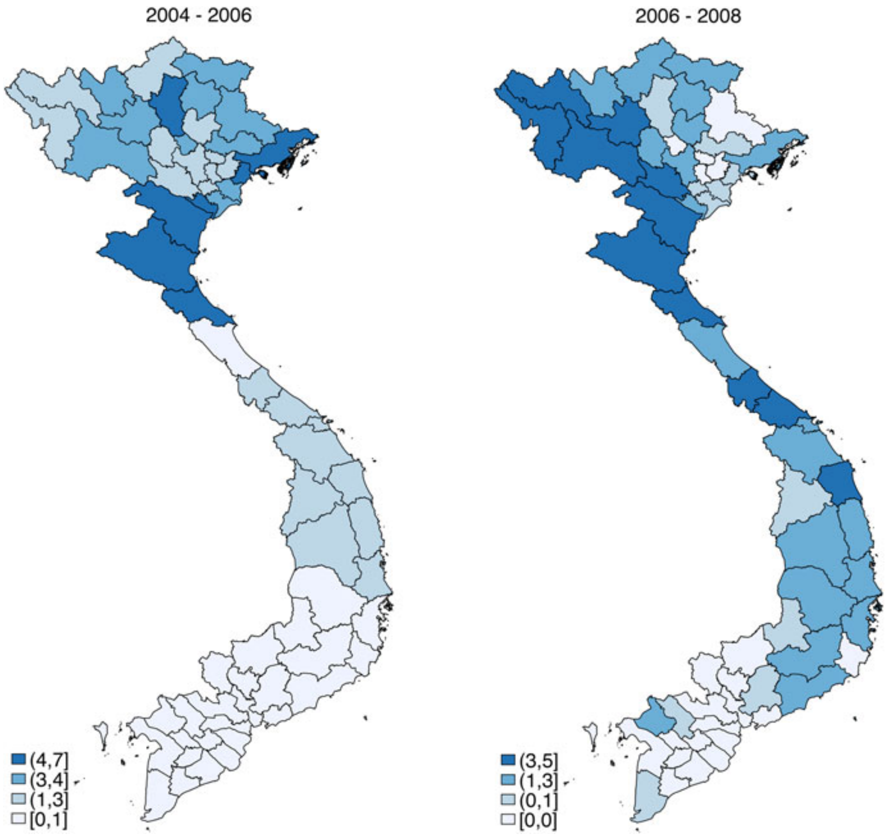

Figure 2. Number of disasters across provinces.
Source: Author's calculation using UNISDR database.

$$\text{Migration}_{i,j,t} = \beta_0 + \beta_1 \text{Crop\_revenue}_{i,j,t-1} + \beta_2' X_{i,j,t} + \beta_3 C_{i,j} + \beta_4 E_{i,j,t} + \sigma_{i,j,t} \quad (2)$$

In equation (1), $\text{Crop\_revenue}_{i,j,t-1}$ of household i in province j in period $t-1$ is calculated by taking the total gross crop revenue (or total sales for each crop) less all costs, divided by the area of agricultural land.[5] The values of total sales and total expenditures are converted into constant 2010 prices using the CPI. With respect to household demographics, $X_{i,j,t}$ is a set of variables representing household size as well as the age and education level of the household head. Household characteristics are chosen based on those of previous studies examining migration and climate change in developing countries [Kubik and Maurel (2016), Koubi *et al.* (2016)]. Following Kubik and Maurel (2016), this paper also includes $C_{i,j}$ as the cost of migration proxied by the distance from the district capital of province where household i lives

[5]As an alternative measure of crop revenue, we also use crop production to account for subsistence farmers. Results are consistent with the main results we present (see Table A7, Appendix).

to the nearest province. $E_{i,j,t}$ is a dummy variable capturing previous migration experience. It is equal to one if household i had previous members moving out during the previous period $t-1$. Finally, we include province characteristics. These include the secondary school completion rate, the number of hospitals, and the provincial poverty rate to control for differences in the level of development.

In equation (2), $\text{Migration}_{i,j,t}$ is a binary variable which equals one if household i in province j has at least one migrant in period t. $\text{Shock}_{j,t-1}$ is a vector of variables representing the severity of natural disasters in period $t-1$ in province j. In this study, it is posited that the migration decision in period t will depend on reductions of agricultural production caused by natural disasters in the previous period $t-1$. The reason is that migration is a costly decision which may take time to save for and enact. In other words, a disaster during period $t-1$ reduces agricultural production in period $t-1$, which leads to a decision to migrate in period t.

Four measures of the severity of shocks are included: (i) number of deaths; (ii) injuries; (iii) houses damaged; and (iv) houses destroyed. As discussed in section 3, this represents a departure from most previous studies which examine the *occurrence* of shocks rather than their *intensity* [Koubi *et al.* (2016)]. One may argue that densely populated areas are likely to have higher exposure to shocks. To account for this, the numbers of deaths and injuries are divided by the population of province j, whereas the number of houses damaged and houses destroyed are divided by the total number of households in the province.

As discussed above, one potential limitation of using the IV framework is that an extreme event can directly impact on migration decisions by damaging or destroying household infrastructure as well as reducing their income from a loss of agricultural output. Natural disaster variables, therefore, should also be considered in the second stage of the IV framework. This, in turn, leads to a problem of collinearity as natural disasters are strongly correlated with agricultural output. If both variables are included in the model, their coefficients might have the wrong sign or implausible magnitudes [Greene (2012)]. However, if one of these variables is omitted from the model, it becomes mis-specified and parameter estimates will be subject to omitted variable bias. To account for this, we adopt a residual generated regressor approach. The first stage of our approach models agricultural output (crop revenue) in period $t-1$ as a function of weather-shock outcome variables and other controls, as indicated in equation (3). One advantage of using the residual generated regressor approach is that it allows us to control for idiosyncratic factors that affect agricultural output in the first stage, whereas the traditional IV approach requires the same set of control variables in both stages:

$$\text{Crop\_revenue}_{i,j,t-1} = \gamma_0 + \gamma_1' \text{Shock}_{j,t-1} + \gamma_2' X_{i,j,t-1} + \gamma_3 Z_{i,j} + \beta_4 I_{i,t-1} + \upsilon_{i,j,t-1} \quad (3)$$

where $Z_{i,j}$ is a set of variables representing the quality of arable land. We assume that arable land is a function of soil type and the irrigation technology available to the household. Type of soil is measured using a dummy variable equal to one if the farm-household has rich ferralitic soil, implying low silica and high aluminum and iron contents. Following Schlenker *et al.* (2005) and Kurukulasuriya and Mendelsohn (2008), we include an irrigation binary variable ($I_{i,t-1}$) which equals one if the household has access to an irrigation system and equals zero if they use rain water. In all model specifications, standard errors are clustered at the province level to control for serial correlation in the error terms across households in the same area.[6]

The residual $\upsilon_{i,j,t-1}$ in equation (3) captures the variation in agricultural output that is not explained by weather shocks (and other control variables). This residual is then employed in the second stage model explaining the decision to migrate. If the "transmission" variable (in this case crop revenue) has a negative impact on migration and shocks have a negative impact on crop revenue, this method will provide larger coefficients on the shock variables. Moreover, the shock variables and the residual capturing agricultural output variables will not be highly correlated allowing for the accurate identification of their parameters:

$$\text{Migration}_{i,j,t} = \delta_0 + \delta_1' \text{Shock}_{j,t-1} + \delta_2 \upsilon_{i,j,t-1} + \delta_3' X_{i,j,t} + \delta_4 C_{i,j} + \delta_5 E_{i,j,t} + \lambda_{i,j,t} \quad (4)$$

4.2 Data

This study uses migration data from the VHLSS in 2006 and 2008.[7] There are 9,189 households interviewed in each survey, with 4,138 households interviewed in both rounds. In each survey, households were asked if there is any member in their families that has moved to another province since the last interview.[8] The survey also includes the reason why they left. This paper identifies migration as those households with a member who left for work due to economic factors. Other reasons for migrating, such as education or marriage, are not considered because they are unlikely consequences of weather shocks.[9] It is possible that the entire household might migrate after a disaster. This study does not examine this type of migration because entire households moving to a different location in 2006 would not be captured in the 2008 survey.

Since we focus on farm-households, it is likely that migration of household members may be affected by the seasonality of agricultural production. To address this issue, we control for the month of interview in all model estimations. Figure 3 provides a pie

[6]The number of clusters in our analysis is 34 which may raise concerns of low-statistical power [Angrist and Pischke (2008), MacKinnon and Webb (2017)]. Therefore, we employ a wild bootstrap method to correct for the low number of clusters [Cameron *et al.* (2008), Cameron and Miller (2015)]. It is estimated using the user-written command *boottest* in Stata [Roodman *et al.* (2019)]. We present the results in Table A10 (in the Appendix) which provides the parameter estimates and 95% confidence intervals of our variables of interest. In most cases, we find that clustering the bootstrapped errors provides comparable estimates with those from our main estimation.

[7]VHLSS is a biennial household survey which has been conducted since 2002. It is important to recognize that the survey relies on a rotating panel of households and is not designed for long panel data analysis from more than two rounds. This analysis therefore does not use VHLSS in 2004 since there are very few households who were covered by three surveys. The survey in 2010 is not used since there was a change in unit codes in Vietnam.

[8]It should be noted that this paper examines internal migration rather than international migration. Previous studies have shown that international migration is an impractical response to extreme events because migrants often have fewer resources, as well as legal and institutional impediments to reach new destinations [Brown (2008)].

[9]One potential issue arising from the use of VHLSS when studying migration is that only officially registered members that have resided in the household for at least 6 months are selected for interview. Therefore, unregistered workers without permanent residence status are ignored in the sample [Pincus and Sender (2008)]. Arguably, this should not be a major concern in this analysis since we are interested in farm-households whose members are very likely to be official members of households.

chart for the migration types in our sample. Approximately 8% of households have migrants for economic reasons, whereas the proportion of migration for education and marriage are 8.7% and 5.5%, respectively.[10] The overall sample in this study includes 1,003 farm households from 34 provinces in Vietnam.[11]

Information on other household characteristics is also provided in the VHLSS. Province level variables, to capture differences in economic development and infrastructure, are collected from the Statistical Yearbook of Vietnam. They include the average secondary school completion rate, the number of hospitals, and the incidence of poverty.

Natural disasters data are sourced from the DesInventar, provided by the United Nations Office for Disaster Risk Reduction (UNISDR).[12] This database contains a historical inventory of disasters and their impacts on goods, infrastructure, lives, and social relations at province and national levels. One advantage of the DesInventar database is its focus on disasters at the province level, whereas other widely used databases, such as EM-DAT, are collected at a country level [Soto (2015)].[13] Weather shocks are defined as events or forces of nature that have catastrophic consequences on people. This study uses four measures of disaster intensity discussed above. In this analysis, we focus on disasters occurring during the period 2006–2008.[14] By employing the magnitude of disasters rather than their occurrence, we account for the possibility that shocks with a higher intensity will trigger migration, whereas shocks causing less damage are less likely to lead to migration. Tables 1 and 2 provide the descriptive statistics of the demographic and disaster variables in our sample, respectively. In the empirical model, the natural disaster variables are measured relative to the population of the province in the year prior to the investigation period (to control for the fact that disasters will have a greater impact in highly populated areas).

To examine whether migrants were different from non-migrants in terms of observed characteristics, we conduct a simple *t*-test using the sample in 2006 (before migration occurred) and 2008 (after migration) separately. The results are presented in Table A1 (Appendix) which show that in most cases, the two groups are comparable. Still, we acknowledge that it is not possible to completely rule out differences between two groups in terms of unobserved characteristics.

[10]It is possible that migration for marriage can lead to assistance to remaining household members in the advent of shocks. Therefore, we examine the robustness of our findings by using migration defined as those moving for economic reasons and marriage. The results show little evidence of disaster induced migration when these migrants are included (see Panel C of Table A5, Appendix).

[11]Since information on the destination of migrant is not available in our sample, we are unable to identify rural–urban migration in this study. Still, we believe that most migration will be rural–urban since migrants are often attracted by higher economic opportunities in urban cities. In our analysis, farm-households are identified as such if their main source of income is from agricultural activities.

[12]Available at https://www.desinventar.net/.

[13]Another source of disaster data is the VHLSS survey in which commune leaders were asked if there were any disasters occurring in the past 2 years. We do not use these data in our paper due to the subjectivity of disaster reporting. In addition, the extent of the damage from the reported disasters is not available in the VHLSS.

[14]As a robustness check, we aggregate disasters which occurred during the period 2004–2006 and examine the impact on migration. We find results that are consistent with our main findings (see Table A9, Appendix).

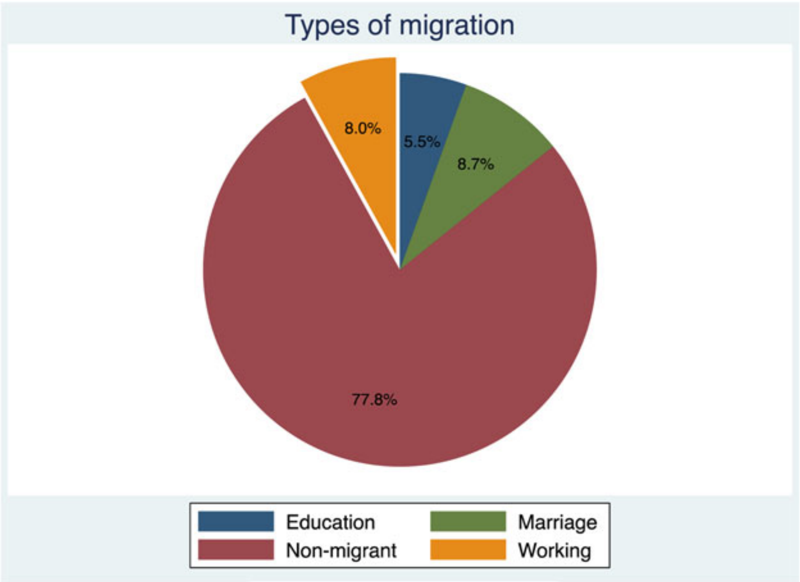

Figure 3. Types of migration.

5. Results

5.1 Main findings

As a starting point, we follow previous studies by employing a traditional IV model using natural disasters as instruments for agricultural production [Cai *et al.* (2016), Kubik and Maurel (2016)]. The results (provided in Table A2 of the Appendix) suggest a negative impact of natural disasters on crop revenue. We examine the potential problem of weak instruments using the Kleibergen–Paap test and compare the test statistics with critical values suggested by Stock and Yogo (2002). The results provide a rejection of the null hypothesis that our instruments are weak. Table A2 also provides the estimation of the second stage of the IV model. It indicates that income loss resulting from weather shocks is associated with a higher probability of migration. However, as noted above, the traditional IV approach does not account for the direct impact of natural disasters on migration. The disaster variables, therefore, should be included in the second stage of the IV model. This, in turn, will lead to the multicollinearity which results in a "wrong" sign or implausible magnitude of the coefficients [Greene (2012)].[15]

We therefore adopt a residual generated regressor approach as presented in equations (3) and (4). Table 3 shows the results from the first stage which examines the impact of natural disasters on agricultural production. It should be noted that the model specification differs from that of the first stage of the traditional IV model, as we are able to control for factors that directly affect agricultural production but not migration decisions, such as soil quality and irrigation. The coefficients on

[15]We also estimate the migration model without controlling for the potential correlation between natural disasters and agricultural production. We then use the variance inflation factor (VIF) to confirm the collinearity problem.

Table 1. Descriptive statistics of demographic variables

Variables	Description	Observations	Mean	Std. Dev.	Min	Max
Farm characteristics						
Crop revenue	Gross crop revenue less all costs, divided by the area of agricultural land	1,003	1,840.050	2,335.434	95	19,500
Ferralitic soil	=1 if ferralitic soil; =0 otherwise	1,003	0.207	0.405	0	1
Irrigated farm	=1 if irrigated farm; =0 if rain-fed farm	1,003	0.627	0.484	0	1
Household head characteristics						
Age	Age of household head (years)	1,003	48.441	12.353	16	91
Education (years)	Education of household head (years)	1,003	6.667	3.471	0	12
Household size	Number of household members	1,003	4.585	1.759	1	14
Province characteristics						
Province area	Total area of province (km^2)	1,003	6,777.133	4,404.510	922	16,499
Population	Total population of province (thous. pers.)	1,003	1,620.189	1,018.576	331	6,347
Secondary completion rate	Rate of secondary school completion at province level (%)	1,003	86.608	7.101	74.240	97.630
Number of hospitals	Number of hospitals at province level	1,003	274.032	176.961	83	685
Poverty rate	Poverty rate at province level (%)	1,003	19.980	9.678	0.5	38
Migration information						
Migration cost (km)	Distance from district capital of province to the nearest province (km)	1,003	3.539	2.109	0.095	7.403
Migration experience	=1 if previous household members are migrants, =0 otherwise	1,003	0.037	0.189	0	1

Table 2. Descriptive statistics of natural disasters by provinces

Variables	Observation	Mean	Std. Dev.	Min	Max
Number of deaths	34	54.522	40.365	0	135
Number of people injured	34	40.743	33.934	0	193
Number of houses destroyed	34	440.271	1,074.644	0	5,770
Number of houses damaged	34	1,486.449	2,064.387	0	7,464

Notes: This table shows the actual number of deaths, injuries, house destroyed, and house damaged. In the empirical model, all variables are normalized by population and number of households.

weather shocks, measured by the number of people injured, houses damaged, and houses destroyed, are negative and statistically significant. All disaster variables are adjusted for population and the number of households. Therefore, we interpret our results in the following way. For the average household in the period 2006–2008, a 10% increase in the number of injuries, houses damaged, and houses destroyed, decreases crop revenue by 0.75%, 0.73%, and 0.79%, respectively.[16] This impact is explained by an unfavorable climatic event during the planting season potentially wiping out arable land and destroying other inputs such as capital equipment and water supply infrastructure. During the non-planting season, impacts of these extreme events might include soil erosion, diseases, and insect infestations.

Table 3 also reveals that many of the coefficients on the control variables are statistically significant. Farmers who use irrigation systems to cultivate crops have higher productivity than those who use rain water. In relation to household characteristics, the coefficients attached to the education variables are statistically significant. Those who have a higher level of education are likely to have better knowledge of cultivation techniques, thereby increasing their agricultural productivity. Similarly, the coefficients on household size have a positive relationship with crop revenue. A possible explanation is that when a household has a higher number of family members, it faces fewer constraints with respect to labor supply.

Table 4 presents the results from the second stage estimation. We use the residual from the first stage as a proxy for agricultural output not explained by weather shocks. This approach reduces the collinearity between the shock and agricultural output variables in the second stage. The positive coefficient estimates on the natural disaster variables confirm the validity of the direct impact of disasters on migration.[17] For example, column 2 suggests that a 10% increase in the severity of natural disasters, measured by the number of people injured, is associated with an increase in the probability of a household sending a migrant by 0.24 points. Similarly, columns 3 and 4 indicate that when weather shocks are measured by the number of houses damaged and houses destroyed, a 10% increase in the intensity of shocks will increase the probability of migration by 0.12 and 0.13 points,

[16]Since most farmers in Vietnam cultivate different crops rather than a single crop, this study aggregates the yields of all crops. Results are consistent when rice crop revenue (the main crop in Vietnam) is used as the dependent variable.

[17]As a robustness check, we examine the impact of shocks on the number of migrants. The ordered probit model is employed to account for the dependent variable being ordinal. The results are presented in Table A6 (Appendix) and are consistent with our main findings.

Table 3. Natural disasters and migration: residual generated regressor approach (first stage)

Dependent variable: log of crop revenue	Natural disasters are measured by			
	Number of deaths	Number of people injured	Number of houses damaged	Number of houses destroyed
Natural disasters	−0.002	−0.075**	−0.073**	−0.079**
	(0.045)	(0.033)	(0.028)	(0.031)
Ferralitic soil	0.112	0.122	0.102	0.101
	(0.109)	(0.109)	(0.099)	(0.098)
Irrigated farm	0.378***	0.383***	0.372***	0.370***
	(0.117)	(0.114)	(0.113)	(0.113)
Age	−0.054***	−0.054***	−0.053***	−0.053***
	(0.013)	(0.012)	(0.013)	(0.013)
Age-squared	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Education	0.005***	0.006***	0.009***	0.009***
	(0.001)	(0.002)	(0.002)	(0.002)
Household size	0.095***	0.103***	0.100***	0.100***
	(0.017)	(0.016)	(0.016)	(0.016)
Constant	4.910***	4.636***	4.527***	4.545***
	(0.403)	(0.422)	(0.462)	(0.462)
Observations	1,003	1,003	1,003	1,003

Notes: Standard errors in parentheses. Standard errors are clustered at the province level. Natural disaster variables are scaled by population and number of households. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

respectively.[18] Note also that the coefficient on the residual (taken from the first stage regression) is negative and statistically significant. This implies that the higher the value of agricultural output, independent of shocks, the lower the probability of migration.[19]

Turning to the remaining variables, the coefficients attached to the age variables are statistically significant using different measurements of natural disasters. Holding other factors constant, households with an older household-head are more likely to have a

[18]As an alternative way to check the nature of selection into migration, we replicate the main model and exclude control variables. We find consistent results as presented in Table A3 (Appendix).

[19]It is possible that natural disasters are not contained within provincial boundaries, leading to spatial correlation in error terms across provinces. We address this issue by using heteroskedastic and autocorrelation consistent (HAC) standard errors suggested by Conley (1999). It is conducted by using Stata command *spatreg* written by Pisati (2001). The results are presented in Table A4 (Appendix). First, we use Moran's I measure and reject the null hypothesis of zero spatial autocorrelation. We then apply HAC estimation in our preferred identification strategy (residual regressor generated approach). We find consistent direct and indirect impacts of natural disasters on migration. We also note that the coefficients on the natural disaster variables are slightly lower in the HAC estimation compared to our main findings in Table 4.

Table 4. Natural disasters and migration: residual generated regressor approach (second stage)

Dependent variable: probability of migration	Natural disasters are measured by: Number of deaths	Number of people injured	Number of houses damaged	Number of houses destroyed
Natural disasters	0.001	0.024***	0.012**	0.013**
	(0.006)	(0.005)	(0.005)	(0.005)
Residual (stage 1)	−0.025**	−0.022**	−0.020**	−0.020**
	(0.011)	(0.011)	(0.010)	(0.010)
Age	0.032***	0.033***	0.033***	0.033***
	(0.009)	(0.009)	(0.008)	(0.008)
Age-squared	−0.000***	−0.000***	−0.000***	−0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Education	0.001	0.002	0.001	0.001
	(0.003)	(0.003)	(0.002)	(0.002)
Household size	0.014***	0.013***	0.014**	0.014**
	(0.005)	(0.005)	(0.006)	(0.006)
Migration cost	−0.001	−0.008*	0.006	0.006
	(0.004)	(0.005)	(0.005)	(0.005)
Migration experience	0.038	0.037	0.018	0.018
	(0.036)	(0.033)	(0.029)	(0.029)
Population	0.008	0.017	0.009	0.010
	(0.046)	(0.050)	(0.051)	(0.051)
Province area	−0.056	0.090	−0.013	−0.008
	(0.202)	(0.224)	(0.191)	(0.191)
Secondary completion rate	−0.001	−0.001	−0.002	−0.002
	(0.002)	(0.003)	(0.003)	(0.003)
Number of hospitals	0.005	−0.001	0.007	0.007
	(0.019)	(0.020)	(0.017)	(0.017)
Poverty rate	0.001	0.001	0.000	0.000
	(0.002)	(0.002)	(0.002)	(0.002)
Fixed effects	Yes	Yes	Yes	Yes
Observations	1,003	1,003	1,003	1,003

Notes: Standard errors in parentheses. Standard errors are clustered at the province level. Results of probit model. Results presented as average marginal effects. Natural disaster variables are normalized by population and number of households. The residual captures crop revenue not explained by weather shocks in the first stage. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

migrant following a shock.[20] Similarly, household size is positively associated with migration. In other words, households with more members are likely to send out their members possibly to diversify household income. Another determinant of migration is the cost of migration. The larger the distance from the district capital of province where households live to the nearest province, the higher the likely cost of migrating, and the lower the probability of migration. The coefficient on the education of the household head variable is positive but statistically insignificant.[21] Finally, we find little evidence of a role of migration experience on migration decisions.

We are also interested in examining the non-linear impacts of natural disasters on migration. We hypothesize that non-linear impacts of shocks might exist since migration is a costly decision. Although a disaster can increase the probability of migration, an extreme event with a high intensity might prevent migration due to financial constraints. In this case, we should expect an inverted-U relationship between intensity of natural disasters and migration. The results in Table 5, however, provide little evidence of such impact since the coefficients on the quadratic variables of the disaster variables are statistically insignificant.

5.2 Robustness checks

Once we confirm the direct and indirect impacts of natural disasters on migration, we conduct a number of robustness checks using the residual generated regressor approach.[22] In all tests, we use the same set of control variables as the main model as well as accounting for time of interview fixed effects. First, one potential threat to the validity of our estimation is the sample selection bias as we mainly focus on farm households. Therefore, we re-estimate our main model in Table 4 by using a full sample which includes non-agricultural households. Given agriculture is the main activity in rural areas of Vietnam, the number of non-farm households is less than farm households (about 40% of the overall sample). The results, shown in Table 6, confirm the negative impact of weather shocks on migration using this sample.

One question that still remains is to what extent our results are affected by using different types of natural disasters. To examine this issue more thoroughly, we distinguish between (1) hydrological disasters including floods and wet mass movements, and (2) meteorological disasters including storm, hailstorm, and extreme temperature.[23] The results in Table 7 indicate that the impact of weather shocks on migration is consistent across disaster types. However, the evidence of migration induced by a meteorological disaster is statistically significant only at the 10% level.

Furthermore, it is possible that using natural disasters data are subject to reporting/measurement error. For example, Felbermayr and Gröschl (2014) argue that collection of

[20]The turning point of age variables is over 100, whereas the maximum age in our dataset is 91. We thus interpret the effect of age on migration as linear. Using the logarithm form of age provides similar results.

[21]Potential explanations for this finding include: (i) the variable capturing the education level of the household head and not the migrant; and (ii) a relatively high correlation between the secondary completion rate (measured at the provincial level) and the education level of the household head (0.28).

[22]As a falsification test, we examine the impact of natural disasters on migration for education and marriage. The results are presented in Table A5 (Appendix) which show that in most cases, the impact of disasters is not statistically significant. It supports our expectation that migration is mainly induced by economic factors.

[23]Other groups of disasters, such as geophysical disasters, are not considered due to their low incidence in Vietnam.

https://doi.org/10.1017/dem.2020.14 Published online by Cambridge University Press

Table 5. Natural disasters and migration: non-linear impacts of disasters

Dependent variable: probability of migration	Natural disasters are measured by			
	Number of deaths	Number of people injured	Number of houses damaged	Number of houses destroyed
Natural disasters	−0.103	0.316*	0.179*	0.207*
	(0.158)	(0.178)	(0.099)	(0.106)
Natural disasters—quadratic term	0.034	−0.025	−0.012	−0.015
	(0.029)	(0.034)	(0.010)	(0.011)
Residual (stage 1)	−0.131*	−0.147**	−0.158**	−0.159**
	(0.068)	(0.072)	(0.073)	(0.073)
Control variables	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes
Observations	1,003	1,003	1,003	1,003

Notes: Standard errors in parentheses. Standard errors are clustered at the province level. Results of second stage of residual regressor generated approach. Results of first stage are the same as Table 3. Results presented as average marginal effects. Natural disasters are instruments for crop revenue. Natural disaster variables are normalized by population and number of households. Other control variables include age, age-squared, education, household size, migration cost, migration experience, population, area of province, secondary completion rate, and number of hospitals. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

disaster data is mostly based on insurance claims or news stories and not on primary geophysical or meteorological data. Strobl (2012) raises a similar concern that natural disaster events are collected from a number of sources which are often inconsistent, and there appears to be greater reporting of events over time. Therefore, we employ long-term change in climate, measured by deviations of precipitation and temperature to measure shocks, as a robustness check. The monthly precipitation and temperature are derived from the Gridded Monthly Time Series (Version 4.01) dataset (GMTS), developed by the Centre for Climatic Research, University of Delaware. This dataset provides global historical estimates of rainfall and temperature for a grid of 0.5 degree by 0.5 degree of latitude and longitude, where the grid nodes are centered on 0.25 degree. Each grid, thus, covers an area of 50 km^2. We then follow Trinh (2018) and use the administrative map of Vietnam to determine the longitude and latitude for each province. These climatic data are then matched with the household data using the four closest grid points in the GMTS dataset. We define deviation of rainfall/temperature as the difference between the actual rainfall/temperature (averaged over 2006–2008, our study period) and the historical average rainfall/temperature (averaged over 40 years from 1970 to 2014). To make geographical units comparable, we standardize these climatic deviations by dividing by their long-term standard deviation. The results are presented in Table A8 (Appendix). First, we find a significant impact of deviations of rainfall and temperature on crop revenue, and there exists a non-linear impact of rainfall deviation. For example, at low levels, higher rainfall will increase agricultural production, possibly by providing more water supply and reducing the cost of irrigation. There is, however, a turning point at which higher rainfall starts to impede crop production.

Table 6. Natural disasters and migration: full sample and non-farm households

Dependent variable: probability of migration	Full sample				Non-farm households			
	Number of deaths	Number of people injured	Number of houses damaged	Number of houses destroyed	Number of deaths	Number of people injured	Number of houses damaged	Number of houses destroyed
Second stage								
Natural disasters	0.029	0.053***	0.053***	0.034***	0.015	0.047***	0.023**	0.023**
	(0.020)	(0.011)	(0.011)	(0.009)	(0.023)	(0.011)	(0.011)	(0.011)
Residual (stage 1)	−0.064***	−0.053**	−0.054**	−0.053**	−0.040**	−0.039**	−0.036**	−0.036**
	(0.022)	(0.023)	(0.022)	(0.022)	(0.016)	(0.016)	(0.015)	(0.015)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
First stage (dependent variable is log of crop revenue)								
Natural disasters	−0.021	−0.044*	−0.051***	−0.058***	−0.028	−0.072***	−0.075***	−0.089***
	(0.038)	(0.026)	(0.018)	(0.020)	(0.017)	(0.020)	(0.014)	(0.015)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,710	1,710	1,710	1,710	706	706	706	706

Notes: Standard errors in parentheses. Standard errors are clustered at the province level. Results of residual regressor generated approach. Results presented as average marginal effects. Natural disaster variables are normalized by population and number of households. Other control variables include age, age-squared, education, household size, migration cost, migration experience, population, area of province, secondary completion rate, number of hospitals, and poverty rate. Control variables in the first stage are the same as Table 3. For non-farm households, we use income per capital as the proxy for crop revenue. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

Table 7. Natural disasters and migration: hydrological disasters and meteorological disasters

Dependent variable: probability of migration	Hydrological disasters: Number of deaths	Hydrological disasters: Number of people injured	Hydrological disasters: Number of houses damaged	Hydrological disasters: Number of houses destroyed	Meteorological disasters: Number of deaths	Meteorological disasters: Number of people injured	Meteorological disasters: Number of houses damaged	Meteorological disasters: Number of houses destroyed
Second stage								
Natural disasters	0.035*	0.051***	0.022**	0.022**	0.012	0.029*	0.018*	0.018*
	(0.020)	(0.011)	(0.009)	(0.009)	(0.012)	(0.013)	(0.010)	(0.011)
Residual (stage 1)	−0.039**	−0.036**	−0.036**	−0.036**	−-0.039**	−0.031**	−0.033**	−0.033**
	(0.016)	(0.015)	(0.014)	(0.014)	(0.017)	(0.013)	(0.013)	(0.013)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
First stage (dependent variable is log of crop revenue)								
Natural disasters	−0.068	−0.087*	−0.085**	−0.095**	−0.513**	−0.021	−0.015	−0.016
	(0.056)	(0.044)	(0.032)	(0.035)	(0.221)	(0.064)	(0.039)	(0.039)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	564	564	564	564	339	339	339	339

Notes: Standard errors in parentheses. Standard errors are clustered at the province level. Results of residual regressor generated approach. Results presented as average marginal effects. Natural disaster variables are normalized by population and number of households. Other control variables include age, age-squared, education, household size, migration cost, migration experience, population, area of province, secondary completion rate, number of hospitals, and poverty rate. Control variables in the first stage are the same as Table 3. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

Table 8. Natural disasters and migration: poor households and non-poor households

Dependent variable: probability of migration	Natural disasters are measured by: Number of deaths	Number of people injured	Number of houses damaged	Number of houses destroyed
Natural disasters	0.040	0.192***	0.096**	0.101**
	(0.062)	(0.069)	(0.040)	(0.040)
Poor households	0.362	−0.413	0.051	0.049
	(0.494)	(0.620)	(0.396)	(0.368)
Natural disasters × Poor households	−0.133	0.086	−0.036	−0.041
	(0.132)	(0.159)	(0.061)	(0.063)
Residual (stage 1)	−0.169**	−0.168**	−0.134*	−0.134*
	(0.072)	(0.077)	(0.072)	(0.072)
Control variables	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes
Observations	1,003	1,003	1,003	1,003

Notes: Standard errors in parentheses. Standard errors are clustered at the province level. Results of second stage of residual regressor generated approach. Results of first stage are the same as Table 3. Results presented as average marginal effects. Natural disasters are instruments for crop revenue. Natural disaster variables are normalized by population and number of households. Other control variables include age, age-squared, education, household size, migration cost, migration experience, population, area of province, secondary completion rate, and number of hospitals. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

In the second stage of the residual regressor generated approach, the reduction of crop revenue, proxied by the residuals in the first stage, leads to higher probability of migration. However, there is little evidence of the direct impact of climatic deviations on migration. It lends support to the hypothesis that natural disasters, often occurring at large-scale, are more likely to trigger migration, rather than long-term climate change.

Finally, we explore whether the impact of natural disasters on migration is conditional on the specific context related to economic and demographic characteristics. First, we examine impacts for poor households vs. non-poor households. On the one hand, poorer households are more vulnerable to weather shocks and might therefore have a higher probability of migration. On the other hand, poorer families are restricted by a lack of capital and weaker networks which might impede migration [Myers *et al.* (2008)]. We use information from the 2008 survey where households were identified as poor using the international poverty line of PPP $1.25-a-day. In our sample, 19% of households live below the poverty line. The coefficient on the interaction term, as presented in Table 8, is not statistically significant suggesting that the impact of shocks on the probability of migration is no different for poor households *vis-à-vis* non poor households.

We also examine impacts for farm-households in the Mekong Delta vs. those in other regions. This is motivated by two factors. First, the Mekong Delta is the

Table 9. Natural disasters and migration: role of government support

Dependent variable: probability of migration	Natural disasters are measured by			
	Number of deaths	Number of people injured	Number of houses damaged	Number of houses destroyed
Natural disasters	0.026	0.183**	0.090**	0.095**
	(0.064)	(0.072)	(0.043)	(0.043)
Receiving support	−0.018	−0.593	−0.303	−0.262
	(0.580)	(0.644)	(0.470)	(0.434)
Natural disasters × Receiving support	0.018	0.182	0.051	0.050
	(0.158)	(0.167)	(0.066)	(0.067)
Residual (stage 1)	−0.192***	−0.183**	−0.144**	−0.144**
	(0.069)	(0.072)	(0.069)	(0.069)
Control variables	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes
Observations	1,003	1,003	1,003	1,003

Notes: Standard errors in parentheses. Standard errors are clustered at the province level. Results of second stage of residual regressor generated approach. Results of first stage are the same as Table 3. Results presented as average marginal effects. Natural disasters are instruments for crop revenue. Natural disaster variables are normalized by population and number of households. Other control variables include age, age-squared, education, household size, migration cost, migration experience, population, area of province, secondary completion rate, number of hospitals, and poverty rate. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

dominant agricultural region in Vietnam (popularly known as the "Rice Bowl" of Vietnam), which accounts for more than half of the country's total rice production. Second, due to the impacts of natural disasters, the government, along with international organizations such as World Bank, has implemented a "Community-based disaster risk management program." The aim of the program is to improve the ability of households to cope with disasters in the region by improving infrastructure and providing support to affected households. Therefore, those living in Mekong Delta might have a lower probability of migration induced by natural disasters. In this analysis, we interact households receiving benefits from the program with the natural disaster variables. The proportion of households participating in the program in our sample is 25%. The results presented in Table 9, however, provide no evidence that households receiving support have a lower probability of migration in the event of natural disasters. This implies that these programs might not have been successful in strengthening resilience to natural disasters.[24]

[24]Still, this finding should be interpreted with caution since it is based on a small sample and households receiving benefits from the program might not be comparable to those who did not receive the support. Further studies should examine the effectiveness of the program.

6. Conclusion

This study provides empirical evidence of disaster and weather shock-induced migration in Vietnam. It contributes to the literature by examining the mechanisms through which natural disasters can influence the decision to migrate. Although previous studies have examined agricultural production as the mechanism through which natural disasters lead to migration, the direct impact of shocks has not been explored in the literature. Using a novel econometric technique, our results confirm the existence of a direct, as well as indirect, impact of natural disasters on migration. Our approach accounts for the correlation between crop revenue and weather shocks which provides more accurate estimations of their parameters. We also depart from much of the existing literature by measuring the damage of natural disasters using different indicators: number of deaths, injuries, houses damaged and houses destroyed. We believe that these indicators provide additional insights into the impact of shocks to findings based on examining their occurrence.

The findings suggest that weather shock-driven migration should be an important policy consideration in Vietnam. With rapid urbanization and development disparities across the country, large cities are likely to face increasing pressure from intensified migration flows after disasters. Since farmers may decide to migrate because of reduced agricultural profit, it is necessary for local governments to improve their awareness and readiness to cope with extreme events. For example, it is essential to improve the provision of infrastructure with a long-term vision including the construction of hazard maps, production plans, organizing volunteer teams to help farmers in the post-disaster period, as well as community training and capacity building.

Our finding is consistent across different groups of household characteristics and geographical characteristics. Still, we acknowledge the limitations of the data we use for the analysis. The information on natural disasters used in our study is available at the province level, which does not allow us to examine migration within a province. Further research examining migration in Vietnam, should seek additional data on different types of population mobility as well as the destinations of migrants to allow a more meaningful assessment of disaster induced migration over time.

Acknowledgements. We would like to thank three anonymous referees for their constructive comments and suggestions that significantly improved the paper.

Conflict of interest. No potential conflict of interest was reported by the authors.

References

Angrist, J. D. and J. S. Pischke (2008) *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, New Jersey: Princeton University Press.

Baez, J., G. Caruso, V. Mueller and C. Niu (2017a) Droughts augment youth migration in Northern Latin America and the Caribbean. *Climatic Change* 140(3–4), 423–435.

Baez, J., G. Caruso, V. Mueller and C. Niu (2017b) Heat exposure and youth migration in Central America and the Caribbean. *American Economic Review* 107(5), 446–450.

Baez, J. E., L. Lucchetti, M. E. Genoni and M. Salazar (2017c) Gone with the storm: rainfall shocks and household wellbeing in Guatemala. *Journal of Development Studies* 53(8), 1253–1271.

Barrios, S., L. Bertinelli and E. Strobl (2006) Climatic change and rural–urban migration: the case of sub-Saharan Africa. *Journal of Urban Economics* 60(3), 357–371.

Beine, M. and C. Parsons (2015) Climatic factors as determinants of international migration. *The Scandinavian Journal of Economics* 117(2), 723–767.

Beine, M. and C. Parsons (2017) Climatic factors as determinants of international migration: Redux. *CESifo Economic Studies* 63(4), 386–402.

Berlemann, M. and M. Steinhardt (2017) Climate change, natural disasters, and migration—a survey of the empirical evidence. *CESifo Economic Studies* 63(4), 353–385.

Bohra-Mishra, P., M. Oppenheimer, R. Cai, S. Feng and R. Licker (2017) Climate variability and migration in the Philippines. *Population and Environment* 38(3), 286–308.

Bohra-Mishra, P., M. Oppenheimer and S. M. Hsiang (2014) Nonlinear permanent migration response to climatic variations but minimal response to disasters. *Proceedings of the National Academy of Sciences* 111(27), 9780–9785.

Brown, O. (2008) The numbers game. *Forced Migration Review* 31(8), 8–9.

Cai, R., S. Feng, M. Oppenheimer and M. Pytlikova (2016) Climate variability and international migration: the importance of the agricultural linkage. *Journal of Environmental Economics and Management* 79, 135–151.

Cameron, A. C., J. B. Gelbach and D. L. Miller (2008) Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics* 90(3), 414–427.

Cameron, A. C. and D. L. Miller (2015) A practitioner's guide to cluster-robust inference. *Journal of Human Resources* 50(2), 317–372.

Cattaneo, C. and G. Peri (2016) The migration response to increasing temperatures. *Journal of Development Economics* 122, 127–146.

Clemens, M. A. (2014) *Does Development Reduce Migration. International Handbook on Migration and Economic Development*. Northampton, Massachusetts: Edward Elgar Publishing.

Clemens, M. A. and H. M. Postel (2018) Deterring emigration with foreign aid: an overview of evidence from low-income countries. *Population and Development Review* 44(4), 667–693.

Conley, T. G. (1999) GMM estimation with cross sectional dependence. *Journal of Econometrics* 92(1), 1–45.

Dallmann, I. and K. Millock (2017) Climate variability and inter-state migration in India. *CESifo Economic Studies* 63(4), 560–594.

Datar, A., J. Liu, S. Linnemayr and C. Stecher (2013) The impact of natural disasters on child health and investments in rural India. *Social Science & Medicine* 76, 83–91.

Drabo, A. and L. M. Mbaye (2015) Natural disasters, migration and education: an empirical analysis in developing countries. *Environment and Development Economics* 20(6), 767–796.

Édes, B. W., F. Gemenne, J. Hill and D. Reckien (2012) *Addressing Climate Change and Migration in Asia and the Pacific: Open Access E-Book*. Mandaluyong City, Philippines: Asian Development Bank.

Ezra, M. and G. E. Kiros (2001) Rural out-migration in the drought prone areas of Ethiopia: a multilevel analysis. *International Migration Review* 35(3), 749–771.

Felbermayr, G. and J. Gröschl (2014) Naturally negative: the growth effects of natural disasters. *Journal of Development Economics* 111, 92–106.

Feng, S., A. B. Krueger and M. Oppenheimer (2010) Linkages among climate change, crop yields and Mexico–US cross-border migration. *Proceedings of the National Academy of Sciences* 107(32), 14257–14262.

Findley, S. E. (1994) Does drought increase migration? A study of migration from rural Mali during the 1983–1985 drought. *International Migration Review* 28(3), 539–553.

Gomanee, K., S. Girma and O. Morrissey (2005) Aid and growth in sub-Saharan Africa: accounting for transmission mechanisms. *Journal of International Development* 17(8), 1055–1075.

Gray, C. L. and V. Mueller (2012) Natural disasters and population mobility in Bangladesh. *Proceedings of the National Academy of Sciences* 109(16), 6000–6005.

Greene, W. H. (2012) *Econometric Analysis*, Seventh Edition. Upper Saddle River, NJ: Prentice Hall.

Gröger, A. and Y. Zylberberg (2016) Internal labor migration as a shock coping strategy: evidence from a typhoon. *American Economic Journal: Applied Economics* 8(2), 123–153.

Gröschl, J. and T. Steinwachs (2017) Do natural hazards cause international migration?. *CESifo Economic Studies* 63(4), 445–480.

Harris, J. R. and M. P. Todaro (1970) Migration, unemployment and development: a two-sector analysis. *American Economic Review* 60(1), 126–142.

Henry, S., B. Schoumaker and C. Beauchemin (2004) The impact of rainfall on the first out-migration: a multi-level event-history analysis in Burkina Faso. *Population and Environment* 25(5), 423–460.

IMHEN and UNDP (2015) Viet Nam special report on managing the risks of extreme events and disasters to advance climate change adaptation. Vietnam Publishing House of Natural Resources, Environment and Cartography, Hanoi, Vietnam.

ISPONRE (2009) Viet Nam Assessment Report on Climate Change (VARCC). Institute of Strategy and Policy on natural resources and environment, Vietnam.

Koerber, K. (2006) *Migration Patterns and Mover Characteristics from the 2005 ACS Gulf Coast Area Special Products*. Washington, DC: US Census Bureau.

Koubi, V., G. Spilker, L. Schaffer and T. Bernauer (2016) Environmental stressors and migration: evidence from Vietnam. *World Development* 79, 197–210.

Kubik, Z. and M. Maurel (2016) Weather shocks, agricultural production and migration: evidence from Tanzania. *Journal of Development Studies* 52(5), 665–680.

Kurukulasuriya, P. and R. Mendelsohn (2008) Crop switching as a strategy for adapting to climate change. *African Journal of Agricultural and Resource Economics* 2(1), 105–126.

Maccini, S. and D. Yang (2009) Under the weather: health, schooling, and economic consequences of early-life rainfall. *American Economic Review* 99(3), 1006–1026.

MacKinnon, J. G. and M. D. Webb (2017) Wild bootstrap inference for wildly different cluster sizes. *Journal of Applied Econometrics* 32(2), 233–254.

Marchiori, L., J.-F. Maystadt and I. Schumacher (2012) The impact of weather anomalies on migration in sub-Saharan Africa. *Journal of Environmental Economics and Management* 63(3), 355–374.

McLeman, R. and B. Smit (2006) Migration as an adaptation to climate change. *Climatic Change* 76(1–2), 31–53.

Mendelsohn, R., W. D. Nordhaus and D. Shaw (1994) The impact of global warming on agriculture: a Ricardian analysis. *American Economic Review* 84(4), 753–771.

Meze-Hausken, E. (2000) Migration caused by climate change: how vulnerable are people inn dryland areas? *Mitigation and Adaptation Strategies for Global Change* 5(4), 379–406.

Morrow-Jones, H. A. and C. R. Morrow-Jones (1991) Mobility due to natural disaster: theoretical considerations and preliminary analyses. *Disasters* 15(2), 126–132.

Myers, C. A., T. Slack and J. Singelmann (2008) Social vulnerability and migration in the wake of disaster: the case of Hurricanes Katrina and Rita. *Population and Environment* 29(6), 271–291.

Pachauri, R. K., M. R. Allen, V. R. Barros, J. Broome, W. Cramer, R. Christ and P. Dasgupta (2014) Climate change 2014: synthesis report. Contribution of Working Groups I, II and III to the fifth assessment report of the Intergovernmental Panel on Climate Change (IPCC).

Peri, G. and A. Sasahara (2019) The impact of global warming on rural–urban migrations: evidence from global big data (No. w25728). National Bureau of Economic Research.

Pincus, J. and J. Sender (2008) Quantifying poverty in Viet Nam: who counts?. *Journal of Vietnamese Studies* 3(1), 108–150.

Pisati, M. (2001) Sg162: tools for spatial data analysis. *Stata Technical Bulletin* 60, 21–37.

Roodman, D., M. Ø. Nielsen, J. G. MacKinnon and M. D. Webb (2019) Fast and wild: bootstrap inference in Stata using boottest. *The Stata Journal* 19(1), 4–60.

Rosenzweig, M. R. and O. Stark (1989). Consumption smoothing, migration, and marriage: evidence from rural India. *Journal of Political Economy* 97(4), 905–926.

Schlenker, W., W. M. Hanemann and A. C. Fisher (2005) Will US agriculture really benefit from global warming? Accounting for irrigation in the hedonic approach. *American Economic Review* 95(1), 395–406.

Soto, A. (2015) Deriving information on disasters caused by natural hazards from limited data: a Guatemalan case study. *Natural Hazards* 75(1), 71–94.

Stark, O. and D. E. Bloom (1985) The new economics of labor migration. *American Economic Review* 75(2), 173–178.

Stock, J. H. and M. Yogo (2002) Testing for weak instruments in linear IV regression: National Bureau of Economic Research Cambridge, Mass., USA.

Strobl, E. (2012) The economic growth impact of natural disasters in developing countries: evidence from hurricane strikes in the Central American and Caribbean regions. *Journal of Development Economics* 97(1), 130–141.

Thiede, B. C. and C. L. Gray (2017) Heterogeneous climate effects on human migration in Indonesia. *Population and Environment* 39(2), 147–172.

Trinh, T. A. (2018) The impact of climate change on agriculture: findings from households in Vietnam. *Environmental and Resource Economics* 71(4), 897–921.

Appendix

Table A1. Descriptive statistics of demographic variables by migration status

Variables	2006 (before migration) Non-migrants	2006 (before migration) Migrants	2006 (before migration) Mean difference significant at 5%	2008 (after migration) Non-migrants	2008 (after migration) Migrants	2008 (after migration) Mean difference significant at 5%
Farm characteristics						
Crop revenue	1,826.007	1,998.803	0.521	2,036.603	2,102.402	0.841
Ferralitic soil	0.208	0.195	0.788	0.167	0.152	0.741
Irrigated farm	0.626	0.634	0.887	0.693	0.652	0.486
Characteristics of household head						
Age	48.12	52.11	0.005	49.780	52.106	0.136
Education	6.628	7.11	0.229	7.332	7.015	0.438
Household size	4.633	4.037	0.003	4.383	3.894	0.016
Poor household	0.192	0.122	0.12	0.129	0.106	0.590
Characteristics of province						
Province area	6,721.759	7,411.238	0.174	5,898.337	7,397.171	0.008
Population	1,611.5	1,719.694	0.357	1,838.536	1,874.264	0.784
Migration information						
Migration cost	3.564	3.251	0.198	2.835	2.689	0.495
Migration experience	0.034	0.073	0.073	0.043	0.076	0.218
Number of observations	916	87		916	87	

Notes: Mean difference is calculated from a *t*-test or a chi-squared test for binary variables, where H_0 is equality of means. Data are collected in 2006 (before migration occurred) and 2008 (after migration).

https://doi.org/10.1017/dem.2020.14 Published online by Cambridge University Press

Table A2. Natural disasters and migration: traditional IV approach

Dependent variable: probability of migration	Natural disasters are measured by: Number of deaths	Number of people injured	Number of houses damaged	Number of houses destroyed
Second stage				
Log of crop revenue	−0.020**	−0.181**	−0.163**	−0.163**
	(0.387)	(0.322)	(0.326)	(0.327)
Control variables	Yes	Yes	Yes	Yes
First stage (dependent variable is log of crop revenue)				
Natural disasters	−0.061**	−0.059**	−0.053***	−0.055***
	(0.025)	(0.025)	(0.016)	(0.017)
Control variables	Yes	Yes	Yes	Yes
Kleibergen–Paap *F*-stat (H_0: weak IV)	61.2	69.7	66.5	66.0
Observations	1,003	1,003	1,003	1,003

Notes: Standard errors in parentheses. Standard errors are clustered at the province level. Natural disasters are instruments for crop revenue. Natural disaster variables are normalized by population and number of households. Other control variables include age, age-squared, education, household size, migration cost, migration experience, population, area of province, secondary completion rate, number of hospitals, and poverty rate. The critical value of the *F*-test from Stock and Yogo (2002) is 16.38. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

Table A3. Natural disasters and migration: residual generated regressor approach (no control variables)

Dependent variable: probability of migration	Natural disasters are measured by: Number of deaths	Number of people injured	Number of houses damaged	Number of houses destroyed
Second stage				
Natural disaster	0.065	0.182***	0.072***	0.074***
	(0.041)	(0.045)	(0.025)	(0.025)
Residual (stage 1)	−0.165**	−0.146**	−0.116*	−0.116*
	(0.066)	(0.072)	(0.070)	(0.070)
Control variables	No	No	No	No
Fixed effects	Yes	Yes	Yes	Yes
First stage (dependent variable is log of crop revenue)				
Natural disaster	0.065	0.182***	0.072***	0.074***
	(0.041)	(0.045)	(0.025)	(0.025)
Control variables	No	No	No	No
Observations	1,003	1,003	1,003	1,003

Notes: Standard errors in parentheses. Standard errors are clustered at the province level. Results of second stage are presented as average marginal effects. Natural disaster variables are normalized by population and number of households. The residual captures crop revenue not explained by weather shocks in the first stage. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

https://doi.org/10.1017/dem.2020.14 Published online by Cambridge University Press

Table A4. Natural disasters and migration: Conley spatial HAC standard errors

	Natural disasters are measured by			
Dependent variable: probability of migration	Number of deaths	Number of people injured	Number of houses damaged	Number of houses destroyed
Second stage				
Natural disaster	0.015*	0.019**	0.011**	0.012**
	(0.009)	(0.008)	(0.005)	(0.006)
Residual (stage 1)	−0.031***	−0.027**	−0.019*	−0.019*
	(0.012)	(0.012)	(0.011)	(0.011)
Control variables	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes
First stage (dependent variable is log of crop production)				
Natural disasters	−0.068***	−0.134***	−0.127***	−0.140***
	(0.025)	(0.028)	(0.017)	(0.018)
Control variables	Yes	Yes	Yes	Yes
Spatial autocorrelation test (null hypothesis: spatial randomization)				
Moran's I	0.068***	0.060***	0.093***	0.098***
Observations	1,003	1,003	1,003	1,003

Notes: Standard errors in parentheses. Results of residual regressor generated approach adjusted for spatial autocorrelation. It is estimated using Stata command *spatreg* written by Pisati (2001). Natural disaster variables are normalized by population and number of households. The residual captures crop production not explained by weather shocks in the first stage. Other control variables include age, age-squared, education, household size, migration cost, migration experience, population, area of province, secondary completion rate, number of hospitals, and poverty rate. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

Table A5. Natural disasters and migration: other types of migration

Dependent variable: probability of migration	Natural disasters are measured by: Number of deaths	Number of people injured	Number of houses damaged	Number of houses destroyed
Panel A: Education				
Natural disasters	−0.020	−0.057	−0.070	−0.079
	(0.075)	(0.063)	(0.051)	(0.056)
Panel B: Marriage				
Natural disasters	0.046	−0.104*	−0.035	−0.038
	(0.046)	(0.056)	(0.034)	(0.037)
Panel C: Marriage and working				
Natural disasters	0.056	0.025	0.023	0.026
	(0.046)	(0.051)	(0.026)	(0.028)
Control variables	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes
Observations	1,003	1,003	1,003	1,003

Notes: Standard errors in parentheses. Standard errors are clustered at the province level. Results of second stage of residual regressor generated approach. Results of first stage are the same as Table 3. Results presented as average marginal effects. Natural disasters are instruments for crop revenue. Natural disaster variables are normalized by population and number of households. Other control variables include age, age-squared, education, household size, migration cost, migration experience, population, area of province, secondary completion rate, number of hospitals, and poverty rate. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table A6. Natural disasters and migration: number of migrants

Dependent variable: number of migrants	Natural disasters are measured by: Number of deaths	Number of people injured	Number of houses damaged	Number of houses destroyed
Natural disaster	0.042	0.223***	0.102***	0.107***
	(0.053)	(0.053)	(0.039)	(0.039)
Residual (stage 1)	−0.197***	−0.184**	−0.148**	−0.148**
	(0.070)	(0.077)	(0.069)	(0.069)
Control variables	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes
Observations	1,003	1,003	1,003	1,003

Notes: Standard errors in parentheses. Standard errors are clustered at the province level. Results of residual regressor generated approach. Ordered probit model is employed in the second stage. Results of first stage are the same as Table 3. Natural disasters are instruments for crop revenue. Natural disaster variables are normalized by population and number of households. Other control variables include age, age-squared, education, household size, migration cost, migration experience, population, area of province, secondary completion rate, number of hospitals, and poverty rate. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table A7. Natural disasters and migration: crop production

	Natural disasters are measured by			
Dependent variable: probability of migration	Number of deaths	Number of people injured	Number of houses damaged	Number of houses destroyed
Second stage				
Natural disaster	0.001	0.024***	0.012**	0.013**
	(0.006)	(0.005)	(0.005)	(0.005)
Residual (stage 1)	−0.030***	−0.026**	−0.024**	−0.024**
	(0.011)	(0.011)	(0.011)	(0.011)
Control variables	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes
First stage (dependent variable is log of crop production)				
Natural disasters	0.000	−0.067***	−0.056***	−0.061***
	(0.020)	(0.018)	(0.012)	(0.013)
Control variables	Yes	Yes	Yes	Yes
Observations	1,003	1,003	1,003	1,003

Notes: Standard errors in parentheses. Standard errors are clustered at the province level. Results of second stage are presented as average marginal effects. Natural disaster variables are normalized by population and number of households. The residual captures crop production not explained by weather shocks in the first stage. Other control variables include age, age-squared, education, household size, migration cost, migration experience, population, area of province, secondary completion rate, number of hospitals, and poverty rate. Control variables in the first stage are the same as Table 3. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

https://doi.org/10.1017/dem.2020.14 Published online by Cambridge University Press

Table A8. Deviation of rainfall/temperature and migration

	First stage	Second stage
Dependent variable:	Log of crop production	Probability of migration
Rainfall deviation	0.730***	−0.011*
	(0.219)	(0.006)
Rainfall deviation-squared	−0.361***	0.000
	(0.115)	(0.000)
Temperature deviation	0.433**	−1.700
	(0.169)	(1.361)
Temperature deviation-squared	0.040	2.656**
	(0.069)	(1.243)
Residual (stage 1)		−0.249***
		(0.087)
Control variables	Yes	Yes
Fixed effects	Yes	Yes
Observations	1,003	1,003

Notes: Standard errors in parentheses Standard errors are clustered at the province level Results of second stage are presented as average marginal effects Deviations of rainfall and temperature are standardized by dividing by long-term standard deviation. The residual captures crop production not explained by climatic variation in the first stage. Other control variables include age, age-squared, education, household size, migration cost, migration experience, population, area of province, secondary completion rate, number of hospitals, and poverty rate. Control variables in the first stage are the same as Table 3. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

Table A9. Natural disasters and migration: disasters in the period 2004–2006

	Natural disasters are measured by			
Dependent variable: probability of migration	Number of deaths	Number of people injured	Number of houses damaged	Number of houses destroyed
Second stage				
Natural disaster	0.012	0.027***	0.016***	0.017***
	(0.008)	(0.009)	(0.005)	(0.005)
Residual (stage 1)	−0.022**	−0.021*	−0.023**	−0.023**
	(0.010)	(0.011)	(0.010)	(0.010)
Control variables	Yes	Yes	Yes	Yes
Fixed effects	Yes	Yes	Yes	Yes
First stage (dependent variable is log of crop revenue)				
Natural disasters	0.067	−0.100**	−0.065*	−0.070*
	(0.064)	(0.038)	(0.033)	(0.037)
Control variables	Yes	Yes	Yes	Yes
Observations	1,003	1,003	1,003	1,003

Notes: Standard errors in parentheses. Standard errors are clustered at the province level. Results of second stage are presented as average marginal effects. Natural disaster variables are normalized by population and number of households. The residual captures crop revenue not explained by weather shocks in the first stage. Other control variables include age, age-squared, education, household size, migration cost, migration experience, population, area of province, secondary completion rate, number of hospitals, and poverty rate. Control variables in the first stage are the same as Table 3. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

Table A10. Natural disasters and migration: residual generated regressor approach using wild bootstrap method

Dependent variable: probability of migration	Natural disasters are measured by			
	Number of deaths	Number of people injured	Number of houses damaged	Number of houses destroyed
Second stage				
Natural disaster				
Estimate	0.008	0.020	0.011	0.011
p value	0.132	0.000	0.024	0.024
Confidence interval	[−0.000, 0.174]	[0.012, 0.029]	[0.002, 0.019]	[0.002, 0.019]
Residual (stage 1)				
Estimate	−0.025	−0.024	−0.026	−0.026
p value	0.006	0.018	0.016	0.016
Confidence interval	[−0.042, −0.007]	[−0.043, −0.004]	[−0.044, −0.006]	[−0.044, −0.007]
First stage (dependent variable is log of crop revenue)				
Natural disaster				
Estimate	−0.006	−0.096	−0.082	−0.089
p value	0.902	0.004	0.008	0.008
Confidence interval	[−0.105, 0.087]	[−0.164, −0.035]	[−0.134, −0.033]	[−0.146, −0.035]

Cite this article: Trinh T-A, Feeny S, Posso A (2021). The impact of natural disasters on migration: findings from Vietnam. *Journal of Demographic Economics* **87**, 479–510. https://doi.org/10.1017/dem.2020.14